The Intel® Pentium® 4 Processor

A 1.40+ GHz Microprocessor in a 0.18 Micron CMOS Process

Michael Upton
Principal Engineer, Desktop Platforms Group
Intel Corporation

Download presentation from:
http:// www.intel.com/pentium4

Overview

- Intel® NetBurst™ micro-architecture goals
- Key aspects of performance
- Intel NetBurst micro-architecture details

Intel® NetBurst™ Micro-architecture Design Goals

- World class floating-point and multi-media performance
- Micro-architecture scales as process improves
  - Minimize RC dominated paths
- Eliminate x86 instruction decode constraints
- Low latency for common operations
- Dramatically improve memory system

Intel NetBurst Micro-architecture Designed For World Class Performance

Typical Program Characteristics

- Floating-point / multi-media applications
  - Low branch misprediction rate
  - Moderate instruction level parallelism
  - Huge memory bandwidth requirement
- Integer-based applications
  - High branch misprediction rate
  - Limited instruction level parallelism
- Greater execution width helps floating-point
- Higher frequency helps integer and floating-point

Micro-architecture Focus Must be Different to Address Each Type of Application

Contributors to Performance

- Processor performance measured by runtime
- Program runtime =
  Instruction count * CPI * Clock period
- Intel® Pentium® 4 processor focuses on all three
  - Reduce instructions executed
    - Streaming SIMD Extensions 2 (SSE2)
  - Reduce CPI
    - Low latency instructions
    - Execution trace cache
    - Improved branch predictor
  - Reduce clock period
    - Pipeline designed for scalability

Optimize All Contributors to Performance

Intel® NetBurst™ Micro-architecture Innovations

- Hyper pipelined technology
- Enhanced branch prediction
- Execution trace cache
- Rapid execution engine
- Improved memory system

New Technologies to Increase Performance

Hyper Pipelined Technology

Basic P6 Pipeline

1	2	3	4	5	6	7		10
Fetch	Fetch	Decode	Decode	Decode	Rename	ROB Rd	Rdy/Sc	Exec

Intro at
733 MHz
.18µ

Basic NetBurst™ Micro-architectu

1	2	3	4	5	6	7	8	9	10	11	12	13	14
TC Nxt IP		TC Fetch		Drive	Alloc	Rename		Que	Sch	Sch	Sch	Disp	Disp

Intro at
≥ 1.40 GHz
.18µ

Industry Leading Performance and Clock Rate

Enhanced Branch Prediction

- Accurate branch prediction is key to enabling longer pipelines
- Dramatic improvement over P6 branch predictor
  - 8x the size (4 KB)
  - Eliminates 1/3 of the mispredictions
- Proven to be better than *all* known predictors
  - (g-share, hybrid, etc.)

Better Prediction Keeps Pipeline on Track

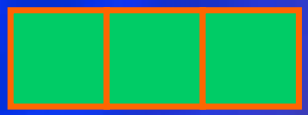
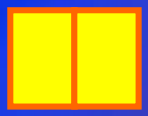
Execution Trace Cache

- Decode of x86 instructions is difficult
- Wide decode can limit frequency
- Trace cache is an advanced Level 1 instruction cache
- Stores 12K decoded micro-ops along predicted path of execution
- Decouples decode from execution path
- Eliminates instruction decode latency
  - Shortens branch misprediction loop

Trace Cache Feeds Fast Engine

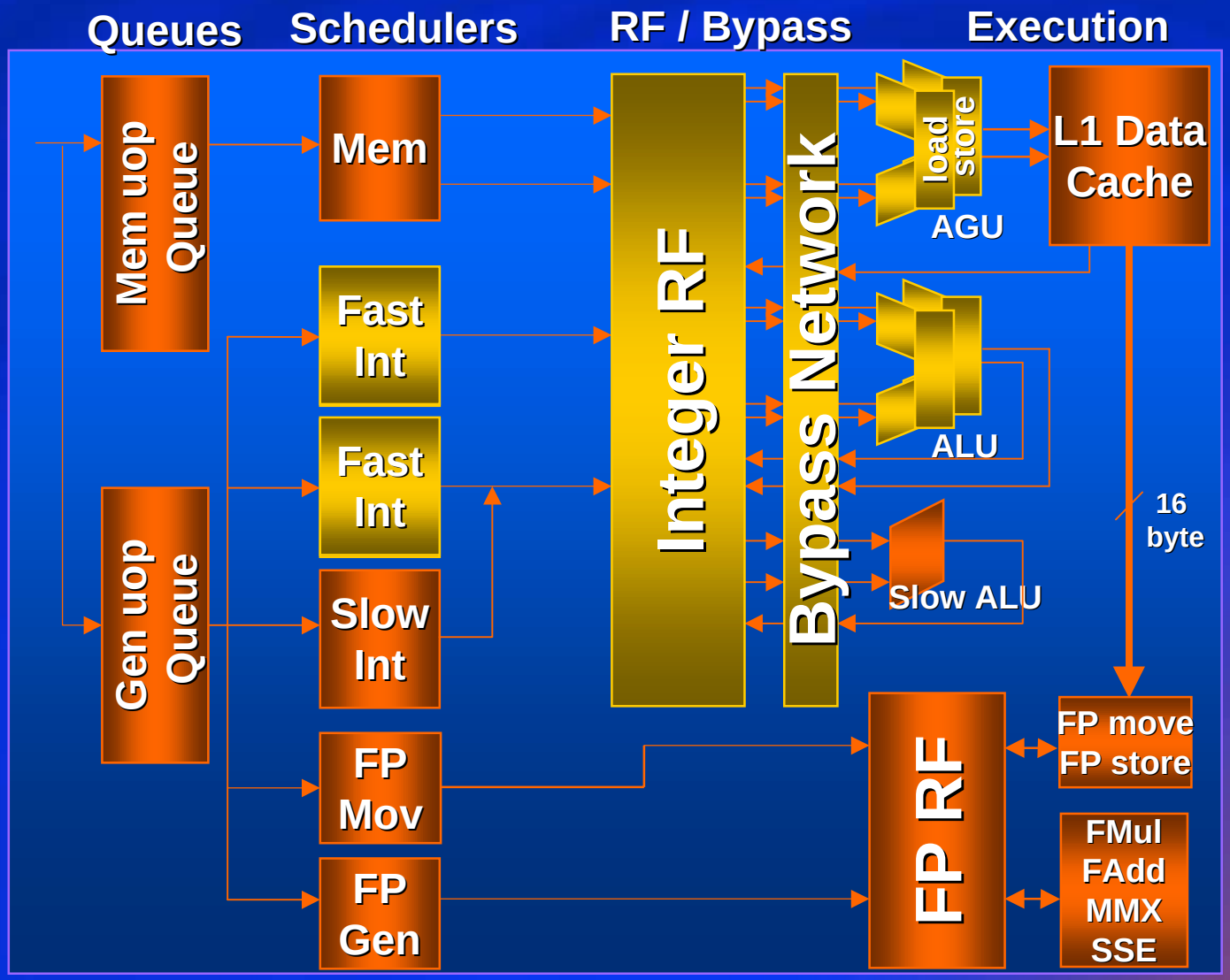
Rapid Execution Engine

- Critical portion of chip runs at 2x clock rate
- Much lower latency for common operations

	ALU	Cache
Intel® Pentium® III Processor @ 1 GHz	 1000 ps	 3000 ps
Intel Pentium 4 Processor @ 1.40 GHz	 360 ps	 1440 ps

Dependent Instructions Processed at High Rate

Core Micro-architecture Details



1.40 GHz

Rapid Execution Engine 2.80 GHz

Data Speculative Execution

- Distance from schedule to execution greater than load latency
- Schedulers dispatch dependent operations before parent has executed
  - Based on cache hit latency
- Track and re-execute instructions that use incorrect data: Replay
- Only dependent operations replay, independent operations can complete

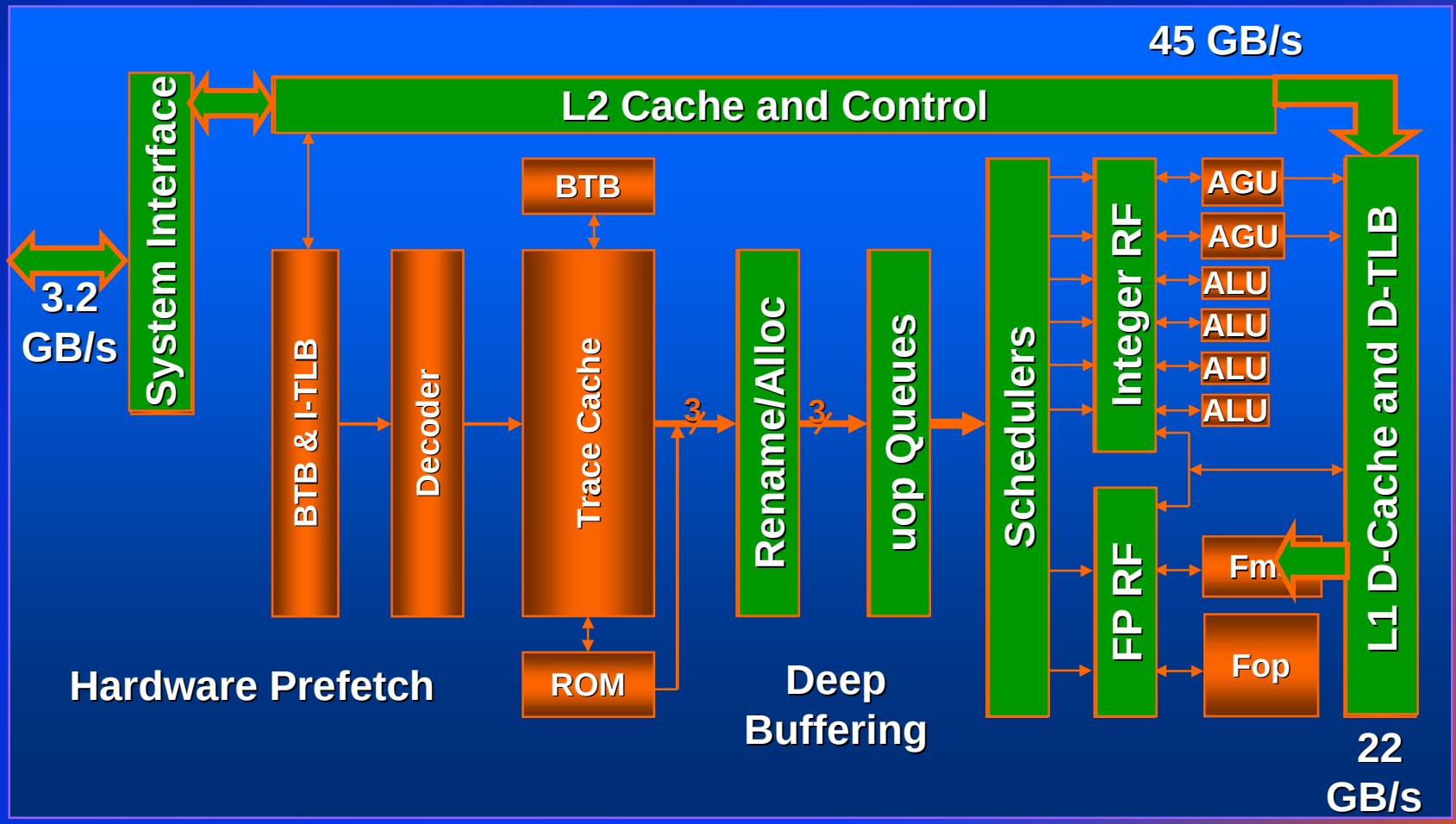
Efficient Mechanism to Reduce Latency

Delivering Floating-Point / Multi-media Performance

- Floating-point performance not set by execution alone
- Two floating-point execution ports
- New SSE2 instructions
- Large instruction window
  - Up to 126 micro-ops in flight
- Focus on memory system
  - High bus bandwidth
  - Long cache lines
  - Hardware prefetch
- Trace cache feeds execution engine

Entire Machine Optimized for Performance

Improved Memory Subsystem



Summary

- Intel® Pentium® 4 processor addresses all three contributors to performance
- Instructions:
  - Streaming SIMD Extensions 2 (SSE2)
- CPI:
  - Low latency instructions
  - Execution trace cache
  - Improved branch predictor
- Frequency:
  - Pipeline designed for scalability

Highly Optimized, Scaleable Micro-architecture for World-class Performance